THE S-R ISSUE: ITS STATUS IN BEHAVIOR ANALYSIS AND IN DONAHOE AND PALMER'S LEARNING AND COMPLEX BEHAVIOR

JOHN W. DONAHOE, DAVID C. PALMER, AND JOSÉ E. BURGOS

UNIVERSITY OF MASSACHUSETTS AT AMHERST, SMITH COLLEGE, AND UNIVERSIDAD CENTRAL DE VENEZUELA AND UNIVERSIDAD CATSLICA DE VENEZUELA

The central focus of this essay is whether the effect of reinforcement is best viewed as the strengthening of responding or the strengthening of the environmental control of responding. We make the argument that adherence to Skinner's goal of achieving a moment-to-moment analysis of behavior compels acceptance of the latter view. Moreover, a thoroughgoing commitment to a moment-to-moment analysis undermines the fundamental distinction between the conditioning processes instantiated by operant and respondent contingencies while buttressing the crucially important differences in their cumulative outcomes. Computer simulations informed by experimental analyses of behavior and neuroscience are used to illustrate these points.

Key words: S-R psychology, contingencies of reinforcement, contiguity, discrimination learning, reinforcement, respondent conditioning, computer simulation

Richard Shull's thoughtful review (Shull, 1995) of Donahoe and Palmer's *Learning and Complex Behavior* (1994) (hereafter, *LCB*) prompted this essay. The review accurately summarized the general themes that informed our efforts and, more to the point for present purposes, identified an important issue—here called the stimulus–response (S-R) issue—that was not directly addressed in our work. Clarifying the status of the S-R issue is important for the further development of behavior analysis, and we seek to make explicit some of the fundamental concerns that surround the issue, most particularly as they arise in LCB.

To provide a context in which to consider the S-R issue, it is helpful to summarize briefly the central themes of the book: (a) Behavior analysis is an independent selectionist science that has a fundamental conceptual kinship with other historical sciences, notably evolutionary biology. (b) Complex behavior, including human behavior, is best understood as the cumulative product of the action over time of relatively simple biobehavioral processes, especially selection by reinforcement. (c) These fundamental processes are characterized through experimental analyses of behavior and, if subbehavioral processes are to be included, of neuroscience. (This contrasts with normative psychology in which subbehavioral processes are inferred from the very behavior they seek to explain, thereby inviting circular reasoning.) (d) Complex human behavior typically occurs under circumstances that preclude experimental analysis. In such cases, understanding is achieved through scientific interpretations that are constrained by experimental analyses of behavior and neuroscience. The most compelling interpretations promise to be those that trace the cumulative effects of reinforcement through formal techniques, such as adaptive neural networks, as a supplement to purely verbal accounts.

It is in the section of the review entitled "Principle of Selection (Reinforcement)" (Shull, 1995, p. 353) that the S-R issue is

raised. The following statement in *LCB* is cited:

> The outcome of selection by reinforcement is a change in the environmental guidance of behavior. That is, what is selected is always an environment–behavior relation, never a response alone. (*LCB*, p. 68)

Of this statement, Shull comments,

> In this respect, then, [*LCB*'s] conception of reinforcement is very much in the tradition of S-R theory . . . [in which] . . . what was selected was the ability of a particular stimulus pattern to evoke a particular response pattern. (Shull, 1995, p. 353)

The question is then considered of whether this view is consistent with the behavior-analytic conception of operant behavior in which "operant behavior occurs in a stimulus context, but there is often no identifiable stimulus change that precedes each occurrence of the response" (Shull, 1995, p. 354). This leads to the related concern of whether adaptive neural networks are suitable to interpret operant behavior because networks are "constructed from elementary connections intended as analogues of stimulus–response relations" (Shull, 1995, p. 354).

In what follows, we seek to demonstrate not only that *LCB*'s view of operant behavior and its interpretation via adaptive neural networks is consistent with behavior-analytic formulations (which we share), but also that this view enriches our understanding of what it means to say that operants are emitted rather than elicited. We agree that the behavior-analytic view of operants should be regarded as "liberating because . . . fundamental relationships could be established in procedures that allowed responses to occur repeatedly over long periods of time without the constraints of trial onset and offset" (Shull, 1995, p. 354). Instead of departing from behavior-analytic thinking, the view that reinforcers select environment–behavior relations fosters more parsimonious treatments of stimulus control and conditioning, and represents a continuation of Skinner's efforts to provide a compelling moment-to-moment account of behavior (Skinner, 1976).

Futher, we concentrate on the rationale behind this view of selection by reinforcement as it is interpreted by biobehaviorally constrained neural networks. Because material relevant to the S-R issue is scattered throughout the book and some of the more technical details are not elaborated, the need for clarification is understandable. We consider first the interpretation of responding in a stable stimulus context and then proceed to a more general examination of the core of the S-R issue. No effort is made to discuss all of its ramifications—the phrase connotes a considerable set of interrelated distinctions that vary somewhat among different theorists (cf. Lieberman, 1993, p. 190; B. Williams, 1986; Zuriff, 1985). Also, no effort is made to provide a historical overview of the S-R issue, although such information is clearly required for a complete treatment of the topic (see Coleman, 1981, 1984; Dinsmoor, 1995; Gormezano & Kehoe, 1981).

BEHAVING IN A STABLE CONTEXT

The central distinction between S-R psychology and the view introduced by Skinner is how one accounts for variability in behavior. The defining feature of S-R psychology is that it explains variability in behavior by reference to variability in antecedents: When a response occurs there must have been some discrete antecedent, or complex of antecedents, overt or covert, that evoked the response. If the response varies in frequency, it is because antecedent events have varied in frequency. On this view, there will always be a nonzero correlation between antecedent events and behavior. Further, frequency of response (or frequency per unit time, i.e., rate) cannot serve as a fundamental dependent variable because response rate is, at root, a function of the rate of stimulus presentation. In contrast, Skinner held that, even when there is no identifiable variability in antecedents, variability in behavior remains lawful: Behavior undergoes orderly change because of its consequences. In fact, at the level of behavioral observations, one can find lawful relationships between the occurrence of a response and the contingencies of reinforcement in a stable context. Skinner did not merely assert the central role of control by consequences; he persuasively demonstrated it experimentally. Once such control is accepted as an empirical fact and not simply as a theoretical preference, the S-R position be-
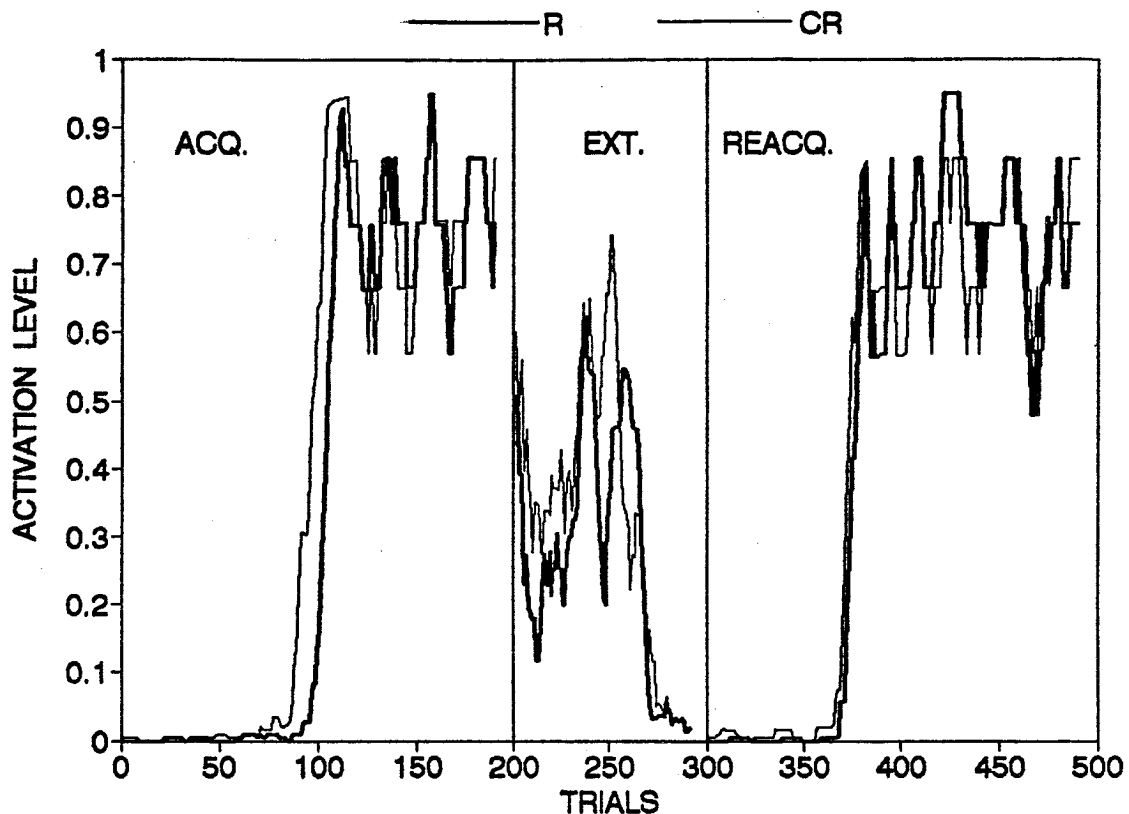
Fig. 1. The simulation by a neural network of acquisition (ACQ), extinction (EXT), and reacquisition (REACQ) with an operant contingency. The simulated environmental context activated the input units of the neural network at a constant level of 1 throughout all phases of the simulation. In accordance with an operant contingency, the input unit for the reinforcing stimulus was activated during ACQ and REACQ only when the activation level of the output unit simulating the operant (R) was greater than zero. During EXT, the input unit for the reinforcing stimulus was never activated. (Activation levels of units could vary between 0 and 1.) The activation level of the output unit simulating the conditioned response (CR), which also changed during the conditioning process, is also shown.

comes untenable. We also accept control by consequences as an empirical fact, and our networks simulate some of its orderly effects without appealing to correlated antecedent changes in the environment.

Consider the neural network simulation of the reacquisition of an extinguished response that is discussed in *LCB* (pp. 92–95). In the first phase of the simulation a response was followed by a reinforcer, in the second phase extinction was scheduled for the response, and in the third phase the response was again reinforced. The "sensory inputs" to the network were held constant throughout the simulation. (Note that in a simulation the stimulus context may be held strictly constant, unaffected by moment-to-moment variations in stimulation that inevitably occur in actual

experiments.) In the simulation, the strength of the response varied widely even though the context remained constant: Responding increased in strength during acquisition, weakened during extinction, and then rapidly increased again during reacquisition, and did so more rapidly than during original acquisition (see Figure 1). Moreover, the changes in response strength were not monotonic, but showed irregularities during the transitions in response strength. None of these changes can be understood by reference to the stimulus context; it remained constant throughout the simulation. Instead, the changes can only be interpreted by reference to the effects of the contingencies of reinforcement on the network and to the history of reinforcement in that context.

BEHAVIORAL AND NEURAL LEVELS OF ANALYSIS

How do we square the foregoing account with the claim that "what is selected is always an environment–behavior relation, never a response alone" (*LCB,* p. 68)? The apparent incongruity arises from a confusion of levels of analysis. We have attempted to uncover relationships between two independent sciences: behavior analysis and neuroscience. Specifically, we have made use of physiological mechanisms that we believe are consistent with behavioral laws. S-R psychology and radical behaviorism are both paradigms of a science of *behavior;* neither includes the underlying physiology in its purview. In a stable context, control by consequences (as opposed to antecedents) stands as a behavioral law, but we propose (at another level of analysis) that the effects of those consequences are implemented by changes in synaptic efficacies. This idea is not new, of course; Watson thought as much (Watson, 1924, p. 209).

Consider how the network accomplishes the simulation discussed above: Changes in the strength of a response occur because of changes in the strengths of connections (simulating changes in synaptic efficacies) along pathways from "upstream" elements. That is, there are changing gradients of control by the constant context as a function of the contingencies of reinforcement. From this perspective, variation in behavior is due to varying consequences, but antecedent events are necessary for the behavior to occur. It is this latter feature of our proposal that encourages the misperception that we are endorsing S-R psychology, because the strength with which an operant unit is activated depends (among other things) on the activation of the inputs of the network by the simulated environment. However, the distinction between S-R psychology and behavior analysis is at the level of behavior, not at the level of biological mechanism. Our networks are intended to simulate aspects of free-operant behavior exhibited by an organism in an experimental chamber and the functioning of the network (i.e., its input–output relations) in conformity with behavioral laws. Thus, we argue that the effects of the consequences of a response are influenced by the context. The analogy with a "black box" is exact: We can eliminate the organism as a variable in our functional relationships, not because the organism is unnecessary, but because it can be ignored in laws of behavior; we treat it as given and as a constant, not as a variable. Similarly, when the context is held constant, it, too, can be ignored, but this does not mean that the context is unnecessary any more than the organism is unnecessary.

In discrimination procedures the context reemerges in our behavioral laws, because it is now a variable. There is a difference between claiming that control by context need not be considered in some situations and claiming that control by context does not exist in those situations. Indeed, Skinner took the first position, not the second (Skinner, 1937). Consider the following: In our simulation of reacquisition, the response gained strength after fewer reinforcers than during original learning because some of the effects of prior reinforcers on the strength of connections within the network had not been completely undone by the intervening period of extinction. The constancy of the context during acquisition and reacquisition played a crucial role in this result because the enduring context permitted some of the same pathways to be activated during both acquisition and reacquisition (cf. *LCB,* p. 94; Kehoe, 1988). With the simulation, as with a living organism, context sets the occasion for responding, although its influence may not be apparent until the context is changed, in which case "generalization decrement" is said to occur. This necessarily implies control by context.

One can interpret observations at the physiological level in ways that more transparently parallel behavioral laws than the accounts we have offered. For example, consider an important finding from Stein's group (e.g., Stein & Belluzzi, 1988, 1989; Stein, Xue, & Belluzzi, 1993, 1994) that is described in *LCB* (p. 56). It was found that the frequency of firing of a neuron could be increased by introducing a neuromodulator, such as dopamine, into the synapse following a burst of firing. These findings have been interpreted to mean that neuromodulators increase the bursting activity of neurons in a manner analogous to the strengthening of emitted behavior by contingent reinforcers. An alternative

interpretation of these same facts will be given that is consistent with our view that reinforcers affect input–output relations and not output alone. However, the primary point here is that it is a mistake to categorize accounts at the behavioral level by one's view of the underlying biology. Behavior does not fully constrain biology. To hold otherwise is to endorse the conceptual-nervous-system approach decried by Skinner (1938).

Consider the following alternative interpretation of the finding that an increase in the frequency of firing occurred as a result of the burst-contingent application of a neuromodulator. The finding was attributed to the contiguity between the bursting of the postsynaptic neuron and the introduction of the neuromodulator (a two-term cellular contingency). Such an interpretation is consistent with the finding, but it is not the only possible interpretation. Moreover, other observations complicate the picture: The neuromodulator was not effective after a single spike, but only after a burst of several spikes. An alternative interpretation of these facts, proposed in *LCB* (pp. 66–67), is that the increase in postsynaptic activity may reflect a heightened sensitivity of the postsynaptic neuron to the release of the neurotransmitter glutamate by presynaptic neurons. The experimental work of Frey (Frey, in press; Frey, Huang, & Kandel, 1993) has shown that dopamine acts in conjunction with the effects of glutamate on the N-methyl-D-aspartate (NMDA) receptor to initiate a second-messenger cascade whose ultimate outcome is an enhanced response of non-NMDA receptors to glutamate. On this view, the ineffectiveness of dopamine after single spikes occurs because bursting is necessary to depolarize the postsynaptic membrane sufficiently to engage the voltage-sensitive NMDA receptor. Accordingly, the increased bursting observed after burst-contingent microinjections of dopamine reflects an enhanced response of the postsynaptic neuron to presynaptic activity (a three-term cellular contingency involving the conjunction of presynaptic and postsynaptic activity with dopamine). To conclude that bursting is independent of presynaptic activity when presynaptic activity has not been measured is to risk mistaking absence of evidence for evidence of absence. In short, interpreting these very important neural observations presents the same conceptual challenge as does interpreting control by a stable context at the behavioral level.

The mechanisms proposed both by Stein and by us are consistent with the behavioral phenomena that led Skinner to break from S-R psychology—an increase in responding following the occurrence of a response-contingent reinforcer in the absence of a specified antecedent. However, we prefer our proposals at both the behavioral and neural levels because they can accommodate behavior in discrimination procedures as well as in stable contexts. It appears to us that proposals that do not specify a three-term contingency must be supplemented by something akin to our proposal in order to account for discriminated behavior, in which case the former proposed mechanisms would be redundant. Ultimately, of course, the interpretation of the cellular results is an empirical matter requiring simultaneous measurement of all terms in the three-term contingency: antecedent events (presynaptic activity), subsequent events (postsynaptic activity), and consequences (the neuromodulator). Both proposals have the merit of showing that behavior analysis can be quite smoothly integrated with what is known about the nervous system. This remains but an elusive dream in normative (i.e., inferred-process) psychology.

In brief, principles formulated on the basis of behavioral observations do not tightly constrain the potential physiological mechanisms that implement the functional relations described by behavioral principles, and physiological mechanisms do not dictate the most effective statement of principles at the behavioral level. The two levels of analysis must yield consistent principles but, as Skinner pointed out (1938, p. 432), nothing that is learned about the physiology of behavior can ever undermine valid behavioral laws.

THE MOMENT-TO-MOMENT CHARACTER OF BIOBEHAVIORAL PROCESSES

Basic to the disposition of the S-R issue is an even more fundamental matter: whether functional relations at the behavioral level are best viewed as emergent products of the outcome of moment-to-moment interactions between the organism and its environment or

whether such regularities are sui generis (i.e., understandable only at the level at which they appear). Skinner clearly favored moment-to-moment analyses (e.g., Ferster & Skinner, 1957). Consider the following statements in "Farewell, my lovely!" in which Skinner (1976) poignantly lamented the decline of cumulative records in the pages of this journal.

> What has happened to experiments where rate changed from moment to moment in interesting ways, where a cumulative record told more in a glance than could be described in a page? . . . [Such records] . . . suggested a really extraordinary degree of control over an individual organism as it lived its life from moment to moment. . . . These "molecular" changes in probability of responding are most immediately relevant to our own daily lives. (Skinner, 1976, p. 218)

Skinner's unwavering commitment to a moment-to-moment analysis of behavior (cf. Skinner, 1983, p. 73) has profound—and underappreciated—implications for the resolution of the S-R issue as well as for other central distinctions in behavior analysis, including the distinction between operant and respondent conditioning itself.

Stimulus Control of Behavior

In *LCB,* an organism is described as "immersed in a continuous succession of environmental stimuli . . . in whose presence a continuous succession of responses . . . is occurring. . . . When a [reinforcing] stimulus is introduced into this stream of events, then . . . selection occurs (cf. Schoenfeld & Farmer, 1970)" (p. 49). At the moment when the reinforcer occurs—what Skinner more casually referred to as "the moment of Truth"— *some* stimulus necessarily precedes the reinforced response in both differential and nondifferential conditioning. That is, at the "moment of reinforcement" (Ferster & Skinner, 1957, pp. 2–3), there is no environmental basis by which to distinguish between the two contingencies. Therefore, no basis exists by which different processes could be initiated for nondifferential as contrasted with differential conditioning (i.e., response strengthening in the first instance and stimulus control of strengthening in the second). If control by contextual stimuli does not occur in nondifferential conditioning, then discrim-

ination becomes an anomaly and requires ad hoc principles that differ from those that accommodate nondifferential conditioning. In such a formulation, the environment would become empowered to control behavior when there were differential consequences, but not otherwise. But, is it credible that reinforcers should strengthen behavior relative to a stimulus with one procedure and not with the other? And, if so, what events present at the "moment of reinforcement" are available to differentiate a reinforced response in a discrimination procedure from a reinforced response in a nondiscrimination procedure? The conclusion that no such events exist led Dinsmoor (1995, p. 52) to make much the same point in citing Skinner's statement that "it is the nature of [operant] behavior that . . . discriminative stimuli are practically inevitable" (Skinner, 1937, p. 273; see also Catania & Keller, 1981, p. 163).

During differential operant conditioning, stimuli are sensed in whose presence a response is followed by a reinforcer. But environment–behavior–reinforcer sequences necessarily occur in a nondiscrimination procedure as well. The two procedures differ with respect to the *reliability* with which particular stimuli are present prior to the reinforced response, but that difference cannot be appreciated on a single occasion. The essence of reliability is repeatability. The distinction emerges as a cumulative product of the occurrence of reinforcers over repeated individual occasions. In laboratory procedures that implement nondifferential conditioning, it is not that no stimuli are sensed prior to the response–reinforcer sequence, but that no stimuli specifiable by the experimenter are *reliably* sensed prior to the sequence.

Conditioning of Behavior

Paradoxically, by strictly parallel reasoning, an acceptance of Skinner's commitment to a moment-to-moment analysis of behavior compels a rejection of a fundamental distinction between the conditioning processes instantiated by respondent and operant procedures. Instead, a moment-to-moment analysis calls for a unified theoretical treatment of the conditioning process, with the environmental control of responding as the cumulative outcome of both procedures.

If an organism is continuously immersed in an environment and is continuously behaving in that environment, then both stimulus and response events necessarily precede and, hence, are potentially affected by the occurrence of a reinforcer regardless of the contingency according to which the reinforcer occurs. In a respondent procedure a specified stimulus, the conditioned stimulus (CS) occurs before the unconditioned stimulus (US). The CS is likely to become a constituent of the selected environment–behavior relation because of its temporal relation to the US. The behavioral constituent of the selected relation includes the response elicited by the US, the unconditioned response (UR). However, because organisms are always behaving, other responses may also precede the US (e.g., orienting responses to the CS; Holland, 1977), although these responses may vary somewhat from moment to moment. As an example of a respondent procedure, if a tone precedes the introduction of food into the mouth, then the tone may continue to guide turning the head toward the source of the tone and come to guide salivating elicited by food. In the operant procedure, the contingency ensures that a specific behavior—the operant—occurs before the reinforcer. Because of its proximity to the reinforcer, the operant is then also likely to become a part of the selected environment–behavior relation. However, because behavior always takes place in an environment, some stimulus must precede the reinforcer although the particular stimulus may vary from moment to moment. For example, a rat may see or touch the lever prior to pressing it and receiving food. From this perspective, respondent and operant conditioning are two different procedural arrangements (i.e., contingencies) that differ with respect to the environmental and behavioral events that are reliably contiguous with the reinforcer. But, this procedural difference need not imply different conditioning processes (*LCB*, pp. 49–50; cf. Donahoe, Burgos, & Palmer, 1993; Donahoe, Crowley, Millard, & Stickney, 1982, pp. 19–23).

The view that reinforcers select environment–behavior relations whatever the procedure and that various procedures differ among themselves in the stimuli and responses that are likely to be present at the moment of selection is consistent with central aspects of Skinner's thinking. As noted in *LCB,*

> Although Skinner's treatment of respondent and operant conditioning emphasized the differences between the two procedures and their outcomes, the present treatment is consistent with his emphasis on the ubiquity of what he called the ''three-term contingency'' (Skinner, 1938, 1953). That is, the reinforcement process always involves three elements—a stimulus, a response, and a reinforcer. There is nothing in a unified treatment of classical and operant conditioning that minimizes the crucially important differences between the outcomes of the two procedures for the interpretation of complex behavior. However, a unified principle does deeply question the view that classical and operant procedures produce two different ''kinds'' of learning or require fundamentally different theoretical treatments. Both procedures select environment–behavior relations but, *because of the differences in the events that reliably occur in the vicinity of the reinforcer, the constituents of the selected relations are different.* (*LCB*, p. 65, emphasis added)

Acknowledging that the organism is always behaving in the presence of *some* environment refines the conceptual treatment of respondents and operants by grounding the distinction on the reliability with which specific stimulus and response events are affected by the two contingencies (cf. Palmer & Donahoe, 1992). On a *single* occasion, there is no basis by which to distinguish a respondent from an operant procedure (cf. Hilgard & Marquis, 1940; Hineline, 1986, p. 63). Others, such as Catania, have appreciated this point:

> It is not clear what differential contingencies could be the basis for discrimination of the contingencies themselves. If we argue that some properties of the contingencies must be learned, to what contingencies can we appeal as the basis for that learning? (Catania & Keller, 1981, p. 163)

The difference in procedures produces crucial differences in their ultimate outcomes, but those different outcomes emerge cumulatively over successive iterations of the same reinforcement process acting in accordance with the specific contiguities instantiated by the procedures. A commitment to a moment-to-moment analysis unavoidably commits one

to the view that reinforcers select environment–behavior relations, not behavior alone. At the "moment of Truth"—whether in a respondent or an operant procedure or in a discrimination or nondiscrimination procedure—the reinforcing stimulus accompanies both environmental and behavioral events. Hence, even if fundamentally different conditioning processes existed for the various procedures, there would be no environmental basis by which one or the other could be appropriately invoked (cf. Donahoe et al., 1982, 1993, pp. 21–22).

In short, we have been misled into searching for different processes to account for respondent and operant conditioning and for nondifferential and differential conditioning, as well as for more complex discrimination procedures (cf. Sidman, 1986), by the language of contingency. *Contingency,* as the term is conventionally used in behavior analysis, refers to relations between events that are defined over repeated instances of the constituent events. We describe our experimental procedures in terms of the manipulation of contingencies, but, by changing the contingencies, we change the contiguities. In our search for the controlling variables, we have confused the experimenter's description of the contingencies with the organism's contact with the contiguities instantiated by those contingencies. And, of course, it is the organism's contact with events, not the experimenter's description of them, that must be the basis for selection by reinforcement.

Contingency is the language of procedure; contiguity is the language of process. We have not thoroughly researched Skinner's use of the term *contingency,* but he employed it, at least sometimes, in a manner that is synonymous with contiguity. For example, "there appears to be no way of preventing the acquisition of non-advantageous behavior through accident. . . . It is only because organisms have reached the point at which a *single* contingency makes a substantial change that they are vulnerable to coincidences" (Skinner, 1953, pp. 86–87, emphasis added; cf. Catania & Keller, 1981, p. 128). (The meaning of contingency as a coincidental relation between events is, in fact, the primary meaning in many dictionaries, although in behavior analysis it much more often denotes reliable relations.)

Relation of Momentary Processes to Molar Regularities

Skinner was resolutely committed to a moment-to-moment account at the behavioral level of analysis, although he did not acknowledge that this view would call for a reassessment of the conceptual distinction between operant and respondent conditioning (but not the crucial differences between these procedures and their corresponding outcomes). His early adherence to a moment-to-moment analysis is apparent in the experimental observation that, under properly controlled circumstances, even a single occurrence of a lever press followed by food changes behavior (Skinner, 1938). Skinner's discussions of superstitious conditioning echo the same theme: Momentary temporal relations may promote conditioning (see also Pear, 1985; Skinner, 1953, pp. 86–87; for alternative interpretations, cf. Staddon & Simmelhag, 1971; Timberlake & Lucas, 1985):

> A stimulus present when a response is reinforced may acquire discriminative control over the response even though its presence at reinforcement is adventitious. (Morse & Skinner, 1957, p. 308)

And,

> to say that a reinforcement is contingent upon a response may mean nothing more than that it follows the response. . . . conditioning takes place because of the temporal relation only, expressed in terms of the order and proximity of response and reinforcement. (Skinner, 1948, p. 168)

The centrality of momentary temporal relations has also been affirmed by students of respondent conditioning. Gormezano and Kehoe, speaking within the associationist tradition, state,

> A single instance of contiguity between A and B may establish an association, repeated instances of contiguity were necessary to establish a cause-effect relation. (p. 3)
> Any relationship of "pairing" or "correlation" can be seen to be an abstraction of the record. (Gormezano & Kehoe, 1981, p. 31)

Moment-to-moment accounts of the conditioning process are also consistent with observations at the neural level. For example, Stein's work indicates that the reinforcing effect of the neuromodulator dopamine occurs

only when it is introduced into the synapse within 200 ms of a burst of firing in the post-synaptic neuron (Stein & Belluzzi, 1989). Behavior analysis and neuroscience are independent disciplines, but their principles cannot be inconsistent with one another's findings. The two sciences are dealing with different aspects of the same organism (*LCB*, pp. 275–277; Skinner, 1938).

Although conditioning processes are instantiated in moment-to-moment relations between events, compelling regularities sometimes appear in the relation between independent and dependent variables defined over more extended periods of time (e.g., between average rate of reinforcement and average rate of responding; Baum, 1973; Herrnstein, 1970). What is the place of molar regularities in a science if its fundamental processes operate on a moment-to-moment basis? Nevin's answer to this question seems very much on the mark: "The possibility that molar relations . . . may prove to be derivative from more local processes does nothing to diminish their value as ways to summarize and integrate data" (Nevin, 1984, p. 431; see also Herrnstein, 1970, p. 253). The conceptual relation between moment-to-moment processes and molar regularities in behavior analysis parallels the distinction between "selection *for*" and "selection *of*" in the paradigmatic selectionist science of evolutionary biology (Sober, 1984). Insofar as the notions of cause and effect have meaning in the context of the complex interchange between an organism and its environment: "'Selection for' describes the causes, while 'selection of' describes the effects" (Sober, 1993, p. 82). In evolutionary biology, selection *for* genes affecting reproductive fitness leads to selection *of* altruistic behavior (Hamilton, 1964). As the distinction applies in behavior analysis, reinforcers cause certain environment–behavior relations to be strengthened; this has the effect, under some circumstances, of producing molar regularities. Selection by reinforcement *for* momentary environment–behavior relations produces selection *of* molar regularities.

One can demonstrate that what reinforcers select are momentary relations between environmental and behavioral events, not the molar regularities that are their cumulative products. This can be done by arranging contingencies of reinforcement that pit moment-to-moment processes against molar regularities. Under these circumstances, the variation in behavior typically tracks moment-to-moment relations, not relations between events defined over more extended periods of time. For example, with positive reinforcers, differential reinforcement of responses that occur at different times following the previous response (i.e., differential reinforcement of interresponse times, or IRTs) changes the overall rate of responding even though the overall rate of reinforcement is unchanged (Platt, 1979). As conjectured by Shimp (1974, p. 498), "there may be no such thing as an asymptotic mean rate of [responding] that is . . . independent of reinforced IRTs" (cf. Anger, 1956). Similarly, in avoidance learning, when the delay between the response and shock is varied but the overall rate of shock is held constant, the rate of avoidance responding is sensitive to the momentary delay between the response and shock, not the overall rate of shock (Hineline, 1970; see also Benedict, 1975; Bolles & Popp, 1964). Research with respondent procedures has led in the same direction: Molar regularities are the cumulative products of moment-to-moment relations. For example, whereas at one time it was held that behavior was sensitive to the overall correlation between conditioned and unconditioned stimuli (Rescorla, 1967), later experiments (Ayres, Benedict, & Witcher, 1975; Benedict & Ayres, 1972; Keller, Ayres, & Mahoney, 1977; cf. Quinsey, 1971) and theoretical work (Rescorla & Wagner, 1972) demonstrated that molar regularities could be understood as the cumulative products of molecular relations between CS and US. In summary, research with both operant and respondent procedures has increasingly shown that molar regularities are the cumulative products of moment-to-moment conditioning processes. (For initial work of this sort, see Neuringer, 1967, and Shimp, 1966, 1969, 1974. For more recent efforts, see Herrnstein, 1982; Herrnstein & Vaughan, 1980; Hinson & Staddon, 1983a, 1983b; Moore, 1984; Silberberg, Hamilton, Ziriax, & Casey, 1978; Silberberg & Ziriax, 1982.)

It must be acknowledged, however, that not all molar regularities can yet be understood as products of molecular processes (e.g., behavior maintained by some schedules or by

long reinforcer delays; Heyman, 1979; Hineline, 1981; Lattal & Gleeson, 1990; Nevin, 1969; B. Williams, 1985). Refractory findings continue to challenge moment-to-moment accounts, and a completely integrated theoretical treatment of molar regularities in terms of molecular processes still eludes us (cf. B. Williams, 1990). Difficulties in providing moment-to-moment accounts of molar regularities in complex situations are not peculiar to behavior analysis. Physics continues to struggle with many-body problems in mechanics, even though all of the relevant fundamental processes are presumably known. Nevertheless, it is now clear that behavior analysis is not forced to choose between molar and moment-to-moment accounts (e.g., Meazzini & Ricci, 1986, p. 37). The two accounts are not inconsistent if the former are regarded as the cumulative product of the latter.

Indeed, the two accounts may be even more intimately intertwined: In the evolutionary history of organisms, natural selection may have favored genes whose expression yielded moment-to-moment processes that implemented certain molar regularities as their cumulative product (*LCB*, pp. 112–114; Donahoe, in press-b; cf. Skinner, 1983, p. 362; Staddon & Hinson, 1983). Natural selection *for* some molar regularity (e.g., maximizing, optimizing, matching) may have led to selection *of* moment-to-moment processes whose product was the molar regularity. In that way, natural selection *for* the molar regularity could lead to selection *of* momentary processes. Once those moment-to-moment processes had been naturally selected, selection by reinforcement *for* momentary environment–behavior relations could, in turn, cause selection *of* the molar regularity. Note, however, to formulate the reinforcement process in terms of the molar regularities it produces, rather than the moment-to-moment processes that implement it, is to conflate natural selection with selection by reinforcement. The selecting effect of the temporally extended environment is the province of natural selection; that of the moment-to-moment environment is the province of selection by reinforcement. Of course, many momentary environments make up the temporally extended environment, but selection by rein-

forcement is *for* the former environments, whereas natural selection is *for* the latter.

Additional experimental work is needed to determine how moment-to-moment processes may lead to molar regularities, but the effort will undoubtedly also require interpretation (Donahoe & Palmer, 1989, 1994, pp. 125–129). In the final section of this essay, interpretation by means of adaptive neural networks is used to clarify the contribution of momentary processes to the central issue: the S-R issue.

NEURAL NETWORK INTERPRETATIONS OF CONDITIONING

We turn finally to the question of whether biobehaviorally constrained neural networks can faithfully interpret salient aspects of the stimulus control of operants. The full answer to this question obviously lies in the future; however, preliminary results are encouraging (e.g., Donahoe et al., 1993; Donahoe & Dorsel, in press; Donahoe & Palmer, 1994). Our concern here is whether, in principle, networks "constructed from elementary connections" that are said to be "analogues of stimulus–response relations" can accommodate the view that "operant behavior occurs in a stimulus context, but there is often no identifiable stimulus change that precedes each occurrence of the response" (Shull, 1995, p. 354). This view of operants is rightly regarded as "liberating" because it empowers the study of complex reinforcement contingencies in the laboratory and because it frees applied behavior analysis from the need to identify the precise controlling stimuli for dysfunctional behavior before instituting remedial interventions. Indeed, it can be argued that pragmatic considerations motivated the operant-respondent distinction more than principled distinctions about the role of the environment in emitted and elicited behavior.

The present inquiry into neural network interpretations of operants can be separated into two parts: The first, and narrower, question is: Do neural networks implement "analogues of stimulus–response relations"? The second is: Are neural networks capable of simulating the effects of nondifferential as well as differential operant contingencies?

Interpreting Environment–Behavior Relations

A neural network consists of (a) a layer of input units whose activation levels simulate the occurrence of environmental events, (b) one or more layers of "hidden" or interior units whose activation levels simulate the states of interneurons, and (c) a layer of output units whose activation levels simulate the effectors that produce behavioral events (cf. Donahoe & Palmer, 1989). If a stimulus–response relation denotes a relation that is mediated by direct connections going from input to output units, then such relations are not, in general, characteristic of neural networks. Although a simple network consisting of only such input–output connections (a so-called perceptron architecture; Rosenblatt, 1962) can mediate a surprising range of input–output relations, some relations that are demonstrable in living organisms are beyond the capabilities of these networks (Minsky & Papert, 1969). In contrast, networks with nonlinear interior units, which more closely simulate the networks of neurons in the nervous system, are typical of modern neural network architectures. Such multilayered networks have already demonstrated their ability to mediate a substantial range of complex environment–behavior relations that are observed with living organisms (e.g., Kehoe, 1988, 1989; cf. McClelland, Rumelhart, & the PDP Research Group, 1986; Rumelhart, McClelland, & the PDP Research Group, 1986). Thus, neither neuroscience nor neural network research endorses formulations in which stimuli guide behavior by means of direct connections akin to monosynaptic reflexes. (We would also note that not even traditional S-R learning theorists—e.g., Guthrie, 1933; Hull, 1934, 1937; Osgood, 1953—held such a simple view of the means whereby the environment guided behavior. In many of their proposals, inferred processes, such as the r_g-s_g mechanism, intervened between the environment and behavior.)

The neural network research of potential interest to behavior analysts is distantly related to what was earlier called S-O-R psychology (where O stood for organism). However, acknowledging a role for the organism in no way endorses an *autonomous* contribution of the organism: All contributions must be traceable to the environment, that is, to histories of selection by the ancestral environment as understood through natural selection and by the individual environment as understood through selection by reinforcement. Also, to be congenial with behavior analysis, all intraorganismic events must be the product of *independent* biobehavioral research; they cannot be inferences from behavior alone. For instance, the organismic counterparts of hidden units are not merely inferences from a behavioral level of observation but are observed entities from a neural level.

In the case of our neural network research, when input units are stimulated by the simulated occurrence of environmental stimuli, the interior units to which those input units are connected are probabilistically activated in the following moment. If a reinforcing signal is present at that moment, then connections are strengthened between input units and all recently activated interior units to which they are connected. The process of strengthening the connections between coactive pre- and postsynaptic units is carried out simultaneously throughout the network at each moment until the end of the simulated time period. The activation levels of units decay over time unless they were reactivated during the preceding moment. Simulations in which the strengths of connections are changed from moment to moment are known as "real-time" simulations, and the successive moments at which the strengths of connections are changed (or "updated") are called "time steps." Stated more generally, real-time neural network simulations implement a dynamical systems approach to the interpretation of behavior (cf. Galbicka, 1992). In a fully realized simulation, the simulated processes that change the strengths of connections, or "connection weights," and the durations of time steps are tightly constrained by independent experimental analyses of neuroscience and behavior (e.g., Buonomano & Merzenich, 1995) and, at a minimum, are consistent with what is known about such processes. Skinner's dictum (1931) that behavior should be understood at the level at which orderly relations emerge applies with equal force to the neural level. Although connection weights are updated on a moment-to-moment basis, the functioning of a network

cannot be understood solely by reference to the environment of the moment: Connection weights at any given moment are a function of the entire selection history of the network to that point. Networks, like organisms, are historic systems whose current performance cannot be understood by reference to the environment of the moment alone (Staddon, 1993; cf. Donahoe, 1993).

Interpreting Behavior in Nondiscrimination Procedures

Before we describe computer simulations that illustrate interpretations of the conditioning of operants, the view that some operants may be uninfluenced by antecedent stimuli requires closer examination. Upon inspection, experimental situations that meet the definition of a nondiscrimination procedure typically contain implicit three-term contingencies. For example, consider a situation in which a pigeon is presented with a response key of a constant green color and key pecking is reinforced with food on some schedule of intermittent reinforcement. Because no other conditions are manipulated by the experimenter, the arrangement is appropriately described as a nondiscrimination procedure. Note, however, that pecking is more likely to be reinforced if the pigeon's head is oriented toward the green key than if it is oriented toward some other stimulus in the situation; pigeons tend to look at what they peck (Jenkins & Sainesbury, 1969). Thus, the observing response of orienting toward the green key is reinforced as a component of a behavioral chain whose terminal response is pecking the green key. Stated more generally, observing responses are often implicitly differentially reinforced in non-discrimination procedures, and the stimuli produced by such responses are therefore more likely to be sensed prior to the reinforced response. As a result, such stimuli come to control the response (Dinsmoor, 1985; cf. Heinemann & Rudolph, 1963).

Moreover, a schedule of reinforcement that is implemented in an environment in which the experimenter has not programmed a relation between features of the environment and the response–reinforcer contingency may nonetheless contain stimuli in whose presence the reinforced response differentially occurs. This relation obtains when non-random environment–behavior relations arise by virtue of the organism's interaction with the environment. And, such interactions generally occur because all responses are not equally likely in the presence of all stimuli. Rats are more apt to make forelimb movements approximating lever pressing (e.g., climbing movements) in environments that contain protruding horizontal surfaces than in environments that are devoid of such features. Behavior is directed toward objects and features of objects, not thin air. When an external environment includes stimuli that make certain behavior more probable, that environment was said to provide "means-end-readinesses" by Tolman (1932) and "affordances" by Gibson (1979).

In addition to stimuli provided by the environment, the organism's own prior behavior produces stimuli that become available to guide further responding. As an example of behaviorally generated stimuli, on ratio schedules a response is more apt to be reinforced following sensory feedback from a burst of prior responses than following feedback from a single prior response (Morse, 1966; D. Williams, 1968). Ferster and Skinner's seminal work, *Schedules of Reinforcement* (1957), is replete with proposals for stimuli that could function as discriminative stimuli in nondiscrimination procedures (see also Blough, 1963; Hinson & Staddon, 1983a, 1983b).

Interpreting Context in Simulations of Operant Conditioning

In the simulation of acquisition, extinction, and reacquisition in a stable environment, the role of context could safely be ignored. However, for reasons noted earlier, control by elements of the context may occur, and that control can be simulated by selection networks, the type of adaptive neural network proposed in *LCB*. Selection networks consist of groups of input units, of interior units simulating neurons in sensory association cortex whose connection strengths are modified by hippocampal efferents, of interior units simulating neurons in motor association cortex whose connection strengths are modified by ventral-tegmental efferents, and of output units.

Figure 2 provides an example of the architecture of a simple selection network (for de-

SENSORY
SUBNETWORK

MOTOR
SUBNETWORK

PRIMARY
SENSORY
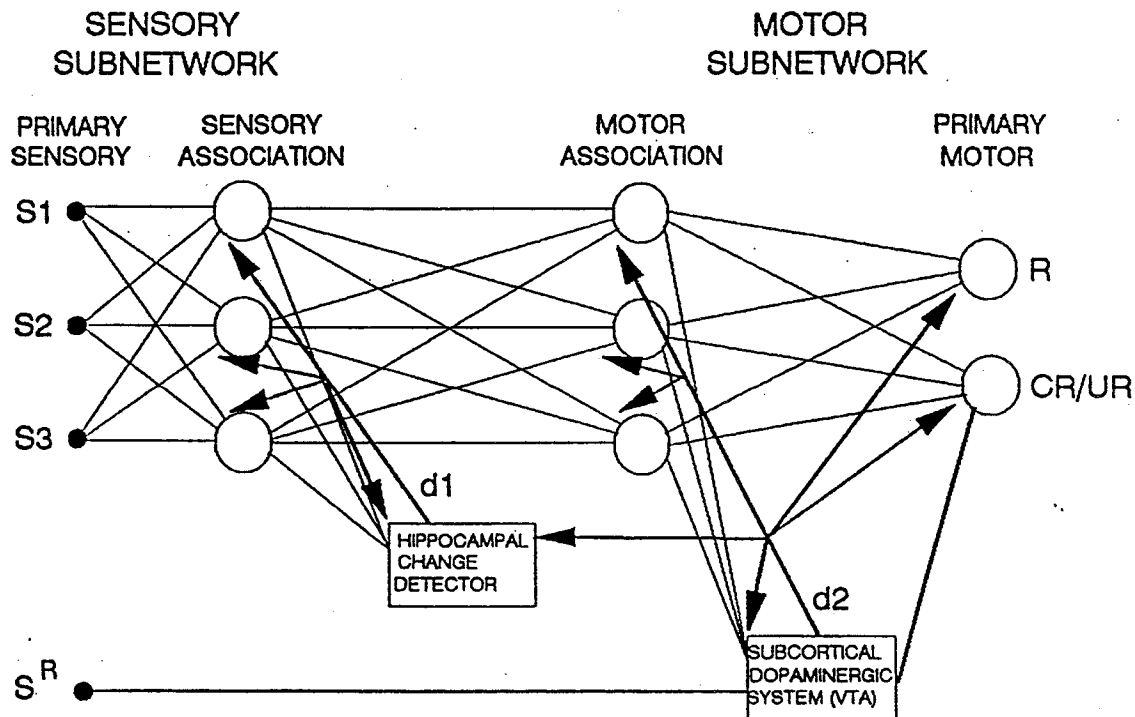
SENSORY
ASSOCIATION

MOTOR
ASSOCIATION

PRIMARY
MOTOR

Fig. 2. A minimal architecture of a selection network for simulating operant conditioning. Environmental events stimulate primary sensory input units (S1, S2, and S3) that give rise to connections that activate units in sensory association areas and, ultimately, units in motor association and primary motor areas. One primary motor output unit simulates the operant response (R). When the R unit is activated, the response–reinforcer contingency implemented by the simulation stimulates the S^R input unit, simulating the reinforcing stimulus. Stimulating the S^R unit activates the subcortical dopaminergic system of the ventral tegmental area (VTA) and the CR/UR output unit simulating the reinforcer-elicited response (i.e., the unconditioned response; UR). Subsequent to conditioning, environmental events acting on the input units permit activation of the R and CR/UR units simulating the operant and conditioned response (CR), respectively. The VTA system modifies connection weights to units in motor association and primary motor areas and modulates the output of the hippocampal system. The output of the hippocampal system modifies connection weights to units in sensory association areas. Connection weights are changed as a function of moment-to-moment changes in (a) the coactivity of pre- and postsynaptic units and (b) the discrepancies in diffusely projecting systems from the hippocampus (d1) and the VTA (d2). The arrowheads point toward those synapses that are affected by activity in the diffusely projecting systems. Finer lines indicate pathways whose connection weights are modified by the diffusely projecting systems. Heavier lines indicate pathways that are functional from the outset of the simulation due to natural selection. (For additional information, see Donahoe et al., 1993; Donahoe & Dorsel, in press; Donahoe & Palmer, 1994.)

tails, see Donahoe et al., 1993; *LCB*, pp. 237–239). A stable context may be simulated using a network with three input units (S1, S2, and S3). In the first simulation, S1 was continuously activated with a strength of .75, simulating a salient feature of the environment (e.g., the wavelength on a key for a pigeon). S2 and S3 were continuously activated with strengths of .50, simulating less salient features of the environment (e.g., the masking noise in the chamber, stimuli from the chamber wall adjacent to the key, etc.). (No simulation can fully capture the complexity and

richness of even the relatively impoverished environment of a test chamber and the relatively simple contingencies programmed therein; Donahoe, in press-a.) Whenever the output unit simulating the operant became activated, a reinforcing stimulus was presented and all connections between recently coactive units were slightly strengthened. After training in which the full context set the occasion for the operant, probe tests were conducted in which each of the three input units making up the context was activated separately and in various combinations. (Note, again,

when even the most salient stimulus, S1, was presented alone and out of context, the operant unit was activated only at a level slightly above .25. As noted in *LCB*, "the environment–behavior relation selected by the reinforcer depends on the context in which the guiding stimulus appears" (p. 139). And, "a stimulus that has been sensed and discriminated may fail to guide behavior when it occurs outside the context in which the discrimination was acquired" (p. 154). The less salient components of the context, S2 and S3, activated the operant unit hardly at all, whether they occurred by themselves or in combination. It was only when S1 was presented in the partial context of either S2 or S3 that the operant unit was strongly activated, although still not as strongly as in the full context.

The lower panel of Figure 3 shows that the effect of context may be even more subtly expressed when no aspect of the context is especially salient. In this simulation, the S1 component of the context was activated at a level of .60 (instead of .75 as in the first simulation), and the S2 and S3 components were activated at a level of .50 as before. Now, when probe tests were simulated, the operant output unit was appreciably activated only by the full context and not by the components, either singly or in combination.

As simulated by selection networks, the environmental guidance of behavior, whether by a specified discriminative stimulus or by components of a variable context, is described in *LCB* as follows:

> Since there are generally a number of possible paths between the relevant input and output units, and since the active pathways mediating the selected input-output relation are likely to vary over time, the selected pathways include a number of alternative paths between the input and output units. Within the network— and that portion of the nervous system the network is intended to simulate—an input unit evokes activity in a class of pathways between the input and output units. At the end of selection, the discriminative [or contextual] stimulus that activates the input units does not so much elicit the response as permit the response to be mediated by one or more of the selected pathways in the network. The . . . stimulus does not elicit the response; it permits the response to be emitted by the organism. (*LCB*, p. 148)
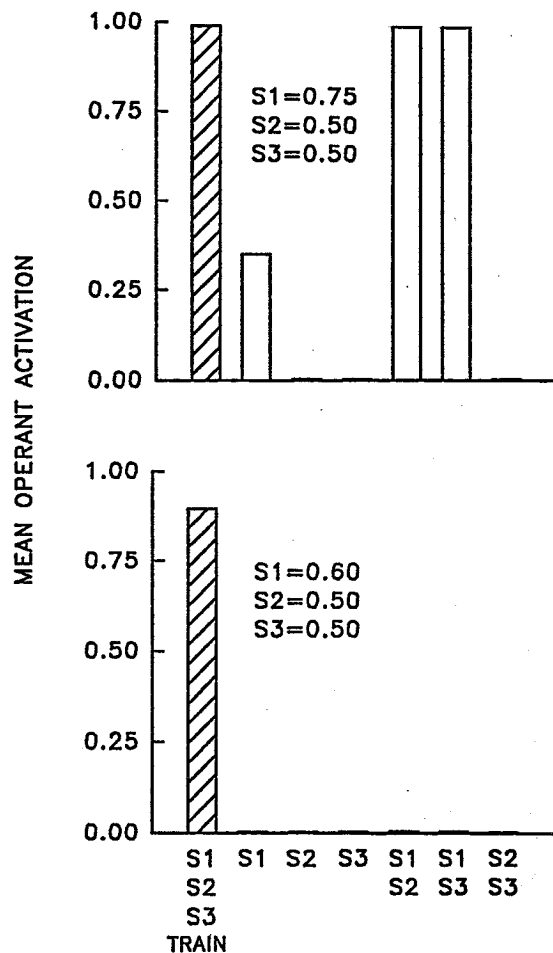
Fig. 3. Simulation results showing the mean activation levels of the operant output unit (R) after conditioning in a stable context consisting of three stimuli (S1, S2, and S3). In the upper panel, S1 was more salient than the other stimuli and activated the S1 input unit at a level of .75 rather than .50 for the S2 and S3 units. In the lower panel, S1 was only slightly more salient than the other stimuli and activated the S1 input unit at a level of .60. The height of each bar represents the mean activation of R by the various stimuli and combinations of stimuli making up the context, including the full context of S1, S2, and S3 used in training (TRAIN).

that simulation permits an assessment of conditions that cannot be completely realized experimentally.)

As shown in the upper panel of Figure 3, by the end of 100 simulated reinforcers following nonzero activation of the operant unit, the output unit was strongly and reliably activated by the full context in which training had taken place (see leftmost bar). However,

On the level of the nervous system, this is the counterpart of Skinner's distinction between elicited responses (respondents) and emitted responses (operants); Skinner, 1937. (*LCB*, p. 151)

Because, in general, behavior is not the result of the environment activating an invariant and rigidly circumscribed set of pathways, *LCB* prefers to speak of behavior as being "guided" rather than controlled by the environment. (As an aside, the phrase "environmental guidance of behavior" has also been found to have certain tactical advantages over "stimulus control of behavior" when seeking a fair hearing for behavior-analytic interpretations of human behavior.)

The foregoing simulations illustrate the context dependence of the conditioning process when an operant is acquired in the stable environment of a nondiscrimination procedure. Our previous simulation research has demonstrated that an operant may be brought under more precise stimulus control: When a discrimination procedure was simulated, the controlling stimuli were restricted to those that most reliably preceded the reinforced response (cf. Donahoe et al., 1993; *LCB*, p. 78). Thus, the same learning algorithm that modifies the strengths of connections in the same selection-network architecture can simulate important conditioning phenomena as its cumulative effect with either a nondiscrimination or a discrimination procedure.

Interpreting the Requirements for Operant Conditioning

Simulation techniques can be applied to the problem of identifying the necessary and sufficient conditions for learning in selection networks. What are the contributions of the stimulus, the two-term response–reinforcer contingency, and the three-term stimulus–response–reinforcer contingency to operant conditioning? And, what role, if any, is played by intranetwork variables that affect the "spontaneous" activity of units?

Consider the question: What is the baseline activation level of the operant unit (i.e., its operant level) when stimuli are applied to input units but without consequences for activity induced in any other units in the network? In living organisms, this condition is imperfectly realized because stimulus presentations by themselves have effects (e.g., habituation, sensitization, or latent inhibition) even when responding has no programmed consequences. However, in a simulation the input units can be stimulated when the algorithms that modify connection weights are disabled. In the present case, when the S1, S2, and S3 input units were stimulated as in the first simulation of context conditioning but with no change in connection weights, the mean activation of the operant output unit during 200 trials was only .09. Thus, stimuli did not evoke activity in the operant unit to any appreciable degree; that is, responding was not elicited.

Turn now to the question: Does conditioning occur if activity of the operant unit is followed by a putative reinforcing stimulus when there is *no* environmental context (not merely no measured or experimenter-manipulated context)? To answer this question, a simulation was conducted under circumstances that were otherwise identical to the first simulation except that the input units of the network were not activated. Any connection strengths that were modified were between units that were activated as the result of spontaneous coactivity between interior and operant units. Under such circumstances, activation of the operant unit is emitted in the purest sense; that is, its activation is solely the product of endogenous intranetwork variables. Simulation indicated that even after as many as 1,000 operant–reinforcer pairings using identical values for all other parameters, conditioning did not occur. Thus, in the absence of an environment, a two-term response–reinforcer contingency was insufficient to produce conditioning in a selection network.

The ineffectiveness of a two-term contingency between an activated output unit and the occurrence of a putative reinforcer is a consequence of our biologically based learning algorithm (Donahoe et al., 1993, p. 40, Equation 5). The learning algorithm simulates modification of synaptic efficacies between neurons, and is informed by experimental analyses of the conditions that produce long-term potentiation (LTP). Experimental analyses of LTP indicate that synaptic efficacies increase when a neuromodulator (that occurs as a result of the reinforcing stimulus) is introduced into syn-
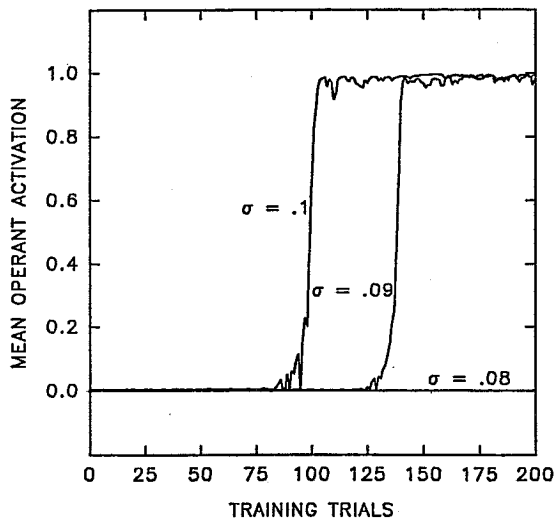
Fig. 4. Simulation results showing changes in the activation level of the R unit during conditioning for different levels of "spontaneous" activity of units in the selection network. The level of spontaneous activity was varied by manipulating the standard deviation (σ) of the logistic function, which determined the activation of a unit as a function of excitation from inputs to that unit. (See text for additional information.)

apses between coactive pre- and postsynaptic neurons (Frey, in press; Frey et al., 1993; see also Beninger, 1983; Hoebel, 1988; Wise, 1989). Under the conditions of the simulation, the presynaptic units and the output unit were very unlikely to be coactive spontaneously. Without stimuli acting on input units to increase the likelihood of coactive units, the simulated reinforcer was ineffective.

Is, then, a three-term contingency sufficient to simulate conditioning in a selection network? The curve in Figure 4 designated by $\sigma = .1$ shows the acquisition function for the first context-conditioning example. After some 75 reinforcers, the operant output unit became increasingly strongly activated. The parameter σ is the standard deviation of the logistic function (see Donahoe et al., 1993, Equation 4), a nonlinear function relating the activation of a postsynaptic unit to the net excitation from its presynaptic inputs. This parameter determines the level of spontaneous activity of a unit. (Neurons in the central nervous system typically have baseline frequencies of firing that are substantially above zero due to local intracellular and extracellular events.)

As shown by the other acquisition functions in Figure 4, reductions in the level of spontaneous activity markedly retarded the simulated acquisition of operant conditioning. With $\sigma = .09$, acquisition did not begin until after 125 reinforcers. Most strikingly, when σ was .08 or less, acquisition failed to occur altogether, even after as many as 200 simulated three-term contingencies. (The level of spontaneous activation of individual units was approximately .001 with $\sigma = .08$.) Thus, in the absence of spontaneous unit activity, even a three-term contingency was insufficient to produce conditioning. From this perspective, the spontaneous activity of neurons is not an impediment to the efficient functioning of the nervous system or to its scientific interpretation by means of neural networks, but is an essential requirement for its operation and understanding.

In conclusion, the effects of a three-term contingency, together with spontaneous unit activity, are necessary and sufficient for the simulation of operant conditioning in selection networks. The interpretation of the selection process by neural networks leads to a deeper understanding of what it means to describe operants as "emitted." In the moment-to-moment account provided by neural networks, the statements that "what is selected is always an environment–behavior relation, never a response alone" (*LCB*, p. 68) and that "operant behavior occurs in a stimulus context, but there is often no identifiable stimulus change that precedes each occurrence of the response" (Shull, 1995, p. 354) are not inconsistent. To the contrary, the statements are complementary: An environment is necessary for reinforcers to select behavior, but without spontaneous intranetwork activity environment–behavior–reinforcer sequences are insufficient. In a moment-to-moment account, as favored by Skinner and implemented by selection networks, environment–behavior relations are neither purely emitted nor purely dependent on particular environment stimuli. Within the range of environment–behavior relations that are conventionally designated as operant, relations are simultaneously guided by the environment *and* emitted by the organism.

REFERENCES

Anger, D. (1956). The dependence of interresponse times upon the relative reinforcement of different in-

terresponse times. *Journal of Experimental Psychology, 52,* 145–161.

Ayres, J. J. B., Benedict, J. O., & Witcher, E. S. (1975). Systematic manipulation of individual events in a truly random control with rats. *Journal of Comparative and Physiological Psychology, 88,* 97–103.

Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior, 20,* 137–154.

Benedict, J. O. (1975). Response-shock delay as a reinforcer in avoidance behavior. *Journal of the Experimental Analysis of Behavior, 24,* 323–332.

Benedict, J. O., & Ayres, J. J. B. (1972). Factors affecting conditioning in the truly random control procedure in the rat. *Journal of Comparative and Physiological Psychology, 78,* 323–330.

Beninger, R. J. (1983). The role of dopamine activity in locomotor activity and learning. *Brain Research Reviews, 6,* 173–196.

Blough, D. S. (1963). Interresponse time as a function of a continuous variable: A new method and some data. *Journal of the Experimental Analysis of Behavior, 6,* 237–246.

Bolles, R. C., & Popp, R. J., Jr. (1964). Parameters affecting the acquisition of Sidman avoidance. *Journal of the Experimental Analysis of Behavior, 7,* 315–321.

Buonomano, D. V., & Merzenich, M. M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science, 267,* 1026–1028.

Catania, A. C., & Keller, K. J. (1981). Contingency, contiguity, correlation, and the concept of causality. In P. Harzem & M. D. Zeiler (Eds.), *Predictability, correlation, and contiguity* (pp. 125–167). New York: Wiley.

Coleman, S. R. (1981). Historical context and systematic functions of the concept of the operant. *Behaviorism, 9,* 207–226.

Coleman, S. (1984). Background and change in B. F. Skinner's metatheory from 1930 to 1938. *Journal of Mind and Behavior, 5,* 471–500.

Dinsmoor, J. A. (1985). The role of observing and attention in establishing stimulus control. *Journal of the Experimental Analysis of Behavior, 43,* 365–381.

Dinsmoor, J. A. (1995). Stimulus control: Part I. *The Behavior Analyst, 18,* 51–68.

Donahoe, J. W. (1993). The unconventional wisdom of B. F. Skinner: The analysis-interpretation distinction. *Journal of the Experimental Analysis of Behavior, 60,* 453–456.

Donahoe, J. W. (in press-a). The necessity of neural networks. In J. W. Donahoe & V. P. Dorsel (Eds.), *Neural-network models of cognition: Biobehavioral foundations.* Amsterdam: Elsevier.

Donahoe, J. W. (in press-b). Positive reinforcement: The selection of behavior. In W. O'Donohue (Ed.), *Learning and behavior therapy.* Boston: Allyn & Bacon.

Donahoe, J. W., Burgos, J. E., & Palmer, D. C. (1993). Selectionist approach to reinforcement. *Journal of the Experimental Analysis of Behavior, 60,* 17–40.

Donahoe, J. W., Crowley, M. A., Millard, W. J., & Stickney, K. A. (1982). A unified principle of reinforcement. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior* (Vol. 2, pp. 493–521). Cambridge, MA: Ballinger.

Donahoe, J. W., & Dorsel, V. P. (Eds.). (in press). *Neural-network models of cognition: Biobehavioral foundations.* Amsterdam: Elsevier.

Donahoe, J. W., & Palmer, D. C. (1989). The interpretation of complex human behavior: Some reactions to *Parallel Distributed Processing. Journal of the Experimental Analysis of Behavior, 51,* 399–416.

Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior.* Boston: Allyn & Bacon.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

Frey, U. (in press). Cellular mechanisms of long-term potentiation: Late maintenance. In J. W. Donahoe & V. P. Dorsel (Eds.), *Neural-network models of cognition: Biobehavioral foundations.* Amsterdam: Elsevier.

Frey, U., Huang, Y.-Y., & Kandel, E. R. (1993). Effects of cAMP simulate a late stage of LTP in hippocampus CA1 neurons. *Science, 260,* 1661–1664.

Galbicka, G. (1992). The dynamics of behavior. *Journal of the Experimental Analysis of Behavior, 57,* 243–248.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton-Mifflin.

Gormezano, I., & Kehoe, E. J. (1981). Classical conditioning and the law of contiguity. In P. Harzem & M. D. Zeiler (Eds.), *Predictability, correlation, and contiguity* (pp. 1–45). New York: Wiley.

Guthrie, E. R. (1933). Association as a function of time interval. *Psychological Review, 40,* 355–367.

Hamilton, W. (1964). The genetical theory of social behavior, I. II. *Journal of Theoretical Biology, 7,* 1–52.

Heinemann, E. G., & Rudolph, R. L. (1963). The effect of discrimination training on the gradient of stimulus generalization. *American Journal of Psychology, 76,* 653–656.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13,* 243–266.

Herrnstein, R. J. (1982). Melioration as behavioral dynamism. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts* (pp. 433–458). Cambridge, MA: Ballinger.

Herrnstein, R. J., & Vaughan, W., Jr. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143–176). New York: Academic Press.

Heyman, G. N. (1979). A Markov model description of changeover probabilities on concurrent variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 31,* 41–51.

Hilgard, E. R., & Marquis, D. G. (1940). *Conditioning and learning.* New York: Appleton-Century-Crofts.

Hineline, P. N. (1970). Negative reinforcement without shock reduction. *Journal of the Experimental Analysis of Behavior, 14,* 259–268.

Hineline, P. N. (1981). The several roles of stimuli in negative reinforcement. In P. Harzem & M. D. Zeiler (Eds.), *Predictability, correlation, and contiguity* (pp. 203–246). New York: Wiley.

Hineline, P. N. (1986). Re-tuning the operant-respondent distinction. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 55–79). Hillsdale, NJ: Erlbaum.

Hinson, J. M., & Staddon, J. E. R. (1983a). Hill-climbing by pigeons. *Journal of the Experimental Analysis of Behavior, 39,* 25–47.

Hinson, J. M., & Staddon, J. E. R. (1983b). Matching,

maximizing, and hillclimbing. *Journal of the Experimental Analysis of Behavior, 40,* 321–331.

Hoebel, B. G. (1988). Neuroscience and motivation: Pathways and peptides that define motivational systems. In R. A. Atkinson (Ed.), *Stevens' handbook of experimental psychology* (Vol. 1, pp. 547–625). New York: Wiley.

Holland, P. C. (1977). Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 77–104.

Hull, C. L. (1934). The concept of habit-family hierarchy and maze learning. *Psychological Review, 41,* 33–54.

Hull, C. L. (1937). Mind, mechanism, and adaptive behavior. *Psychological Review, 44,* 1–32.

Jenkins, H. M., & Sainesbury, R. S. (1969). The development of stimulus control through differential reinforcement. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 123–161). Halifax, Nova Scotia: Dalhousie University Press.

Kehoe, E. J. (1988). A layered network model of associative learning: Learning to learn and configuration. *Psychological Review, 95,* 411–433.

Kehoe, E. J. (1989). Connectionist models of conditioning: A tutorial. *Journal of the Experimental Analysis of Behavior, 52,* 427–440.

Keller, R. J., Ayres, J. J. B., & Mahoney, W. J. (1977). Brief versus extended exposure to truly random control procedures. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 53–65.

Lattal, K. A., & Gleeson, S. (1990). Response acquisition with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 16,* 27–39.

Lieberman, P. A. (1993). *Learning: Behavior and cognition.* Pacific Grove, CA: Brooks/Cole.

McClelland, J. L., Rumelhart, D. E., & The PDP Research Group. (Eds.). (1986). *Parallel distributed processing: Explorations in microstructure of cognition* (Vol. 2). Cambridge, MA: MIT Press.

Meazzini, P., & Ricci, C. (1986). Molar vs. molecular units of analysis. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 19–43). Hillsdale, NJ: Erlbaum.

Minsky, M. L., & Papert, S. A. (1969). *Perceptrons.* Cambridge, MA: MIT Press.

Moore, J. (1984). Choice and transformed interreinforcement intervals. *Journal of the Experimental Analysis of Behavior, 42,* 321–335.

Morse, W. H. (1966). Intermittent reinforcement. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 52–108). New York: Appleton-Century-Crofts.

Morse, W. H., & Skinner, B. F. (1957). A second type of superstition in the pigeon. *American Journal of Psychology, 70,* 308–311.

Neuringer, A. J. (1967). *Choice and rate of responding in the pigeon.* Unpublished doctoral dissertation, Harvard University.

Nevin, J. A. (1969). Interval reinforcement of choice behavior in discrete trials. *Journal of the Experimental Analysis of Behavior, 12,* 875–885.

Nevin, J. A. (1984). Quantitative analysis. *Journal of the Experimental Analysis of Behavior, 42,* 421–434.

Osgood, C. E. (1953). *Method and theory in experimental psychology.* New York: Oxford University Press.

Palmer, D. C., & Donahoe, J. W. (1992). Essentialism and selectionism in cognitive science and behavior analysis. *American Psychologist, 47,* 1344–1358.

Pear, J. J. (1985). Spatiotemporal patterns of behavior produced by variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 44,* 217–231.

Platt, J. R. (1979). Interresponse-time shaping by variable-interval-like interresponse-time reinforcement contingencies. *Journal of the Experimental Analysis of Behavior, 31,* 3–14.

Quinsey, V. L. (1971). Conditioned suppression with no CS-US contingency in the rat. *Canadian Journal of Psychology, 25,* 69–82.

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review, 74,* 71–80.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Rosenblatt, F. (1962). *Principles of neurodynamics.* Washington, DC: Spartan.

Rumelhart, D. E., McClelland, J. L., & The PDP Research Group. (Eds.) (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1). Cambridge, MA: MIT Press.

Schoenfeld, W. N., & Farmer, J. (1970). Reinforcement schedules and the "behavior stream." In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 215–245). New York: Appleton-Century-Crofts.

Shimp, C. P. (1966). Probabilistically reinforced choice behavior in pigeons. *Journal of the Experimental Analysis of Behavior, 9,* 443–455.

Shimp, C. P. (1969). Optimal behavior in free-operant experiments. *Psychological Review, 76,* 97–112.

Shimp, C. P. (1974). Time allocation and response rate. *Journal of the Experimental Analysis of Behavior, 21,* 491–499.

Shull, R. L. (1995). Interpreting cognitive phenomena: Review of Donahoe and Palmer's *Learning and Complex Behavior. Journal of the Experimental Analysis of Behavior, 63,* 347–358.

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213–245). Hillsdale, NJ: Erlbaum.

Silberberg, A., Hamilton, B., Ziriax, J. M., & Casey, J. (1978). The structure of choice. *Journal of Experimental Psychology: Animal Behavior Processes, 4,* 368–398.

Silberberg, A., & Ziriax, J. M. (1982). The interchange-over time as a molecular dependent variable in concurrent schedules. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts of behavior* (pp. 111–130). Cambridge, MA: Ballinger.

Skinner, B. F. (1931). The concept of the reflex in the study of behavior. *Journal of General Psychology, 5,* 427–458.

Skinner, B. F. (1937). Two types of conditioned reflex: A reply to Konorski and Miller. *Journal of General Psychology, 16,* 272–279.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1948). "Superstition" in the pigeon. *Journal of Experimental Psychology, 38,* 168–172.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Skinner, B. F. (1976). Farewell, my lovely! *Journal of the Experimental Analysis of Behavior, 25,* 218.

Skinner, B. F. (1983). *A matter of consequences.* New York: Knopf.

Sober, E. (1984). *The nature of selection.* Cambridge, MA: MIT Press.

Sober, E. (1993). *Philosophy of biology.* Boulder, CO: Westview Press.

Staddon, J. E. R. (1993). The conventional wisdom of behavior analysis. *Journal of the Experimental Analysis of Behavior, 60,* 439–447.

Staddon, J. E. R., & Hinson, J. M. (1983). Optimization: A result or a mechanism? *Science, 221,* 976–977.

Staddon, J. E. R., & Simmelhag, V. L. (1971). The "superstition" experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review, 78,* 3–43.

Stein, L., & Belluzzi, J. D. (1988). Operant conditioning of individual neurons. In M. L. Commons, R. M. Church, J. R. Stellar, & A. R. Wagner (Eds.), *Quantitative analyses of behavior* (Vol. 7, pp. 249–264). Hillsdale, NJ: Erlbaum.

Stein, L., & Belluzzi, J. D. (1989). Cellular investigations of behavioral reinforcement. *Neuroscience and Biobehavioral Reviews, 13,* 69–80.

Stein, L., Xue, B. G., & Belluzzi, J. D. (1993). A cellular analogue of operant conditioning. *Journal of the Experimental Analysis of Behavior, 60,* 41–53.

Stein, L., Xue, B. G., & Belluzzi, J. D. (1994). In vitro reinforcement of hippocampal bursting: A search for Skinner's atom of behavior. *Journal of the Experimental Analysis of Behavior, 61,* 155–168.

Timberlake, W., & Lucas, G. A. (1985). The basis of superstitious behavior: Chance contingency, stimulus substitution, or appetitive behavior? *Journal of the Experimental Analysis of Behavior, 44,* 279–299.

Tolman, E. C. (1932). *Purposive behavior in animals and men.* New York: Appleton-Century-Crofts.

Watson, J. B. (1924). *Behaviorism.* New York: Norton.

Williams, B. A. (1985). Choice behavior in a discrete-trial concurrent VI-VR: A test of maximizing theories of matching. *Learning and Motivation, 16,* 423–443.

Williams, B. A. (1986). Identifying behaviorism's prototype: A review of *Behaviorism: A Conceptual Reconstruction* by G. E. Zuriff. *The Behavior Analyst, 9,* 117–122.

Williams, B. A. (1990). Enduring problems for molecular accounts of operant behavior. *Journal of Experimental Psychology: Animal Behavior Processes, 16,* 213–216.

Williams, D. R. (1968). The structure of response rate. *Journal of the Experimental Analysis of Behavior, 11,* 251–258.

Wise, R. A. (1989). The brain and reward. In J. M. Liebman & S. J. Cooper (Eds.), *The neuropharmacological basis of reward* (pp. 377–424). New York: Oxford University Press.

Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction.* New York: Columbia University Press.